

# Collaborative Grid Services Architecture for Active Thermochemical Tables

Kaizar Amin<sup>1,3,\*</sup>, Ruscic Branko<sup>2</sup>, Gregor von Laszewski<sup>1,†</sup>,  
Sandeep Nijsure<sup>1,3</sup>, Pinzon Reinhardt<sup>2</sup>, Melita Morton<sup>2</sup>,  
Baoshan Wang<sup>2</sup>, Sandra Bittner<sup>1</sup>, Michael Minkoff<sup>1</sup>,  
and Albert Wagner<sup>2</sup>

<sup>1</sup>Mathematics and Computer Science Division, Argonne National Laboratory

<sup>2</sup>Chemistry Division, Argonne National Laboratory

<sup>3</sup>Department of Computer Science, University of North Texas

## Abstract

Based on the requirements of creating the next generation of global scientific collaborative environments we describe our efforts to design and implement a framework called Open Collaborative Grid Services Architecture (OCGSA). Our services enhance the functionality of the Open Grid Services Architecture (OGSA) significantly, by introducing the concept of collaborative groups as part of the overall architectural service design. Hence, our novel services are able to deal with metadata regarding the group creator, name, description, members, and the desired level of event archiving. It also provides the interface to handle security policies at different levels of the application and group, community interaction mechanisms, an adequate presence management for users, and capabilities for interactively controlling and steering the application. We are applying these services to an important collaborative application in the chemical sciences, called Active Thermochemical Tables (ATcT), to create thermochemistry data in a collaborative fashion with the thermochemistry community which will be far more accurate than the data currently available. We expect the data to be shared as part of an international peer reviewed effort to verify, improve, and enhance thermochemistry data across multiple scales.

---

\*amin@mcs.anl.gov

†gregor@mcs.anl.gov

# 1 Introduction

Thermochemistry is one of the essential and undeniably fundamental building blocks of multi-scale chemical science. Knowledge of thermochemical stability of substances is central to almost all aspects of chemistry and critical in many industries, because the outcome of chemical reactions is ultimately governed by thermodynamics. In particular, accurate, reliable, and internally consistent thermochemical values are a *conditio sine qua non* in kinetics, construction of reaction mechanisms, and formulation of descriptive models that have predictive abilities. Thermochemistry also influences the development of theory, where accurate values serve as benchmarks for evaluating and/or developing new electronic structure calculations. In addition, accurate thermochemical values are frequently used as empirical correctors for adjusting calculated potential energy surfaces, etc. Finally, knowledge of accurate thermochemistry is often able to generate new insights on chemical bonding in molecules.

Traditional thermochemistry uses a sequential approach to developing thermochemical values from measurable quantities such as bond dissociation energies, equilibrium constants, reaction enthalpies, etc. The resulting thermochemical values, i.e. solutions obtained by relating these measurable quantities according to thermodynamic rules, are usually presented in tabulations known as thermochemical tables. The traditional approach has very serious disadvantages and limits. For example, the tabulated values have intricate, but hidden, progenitor-progeny dependences. These hidden dependencies make traditional thermodynamic tables *static*, in the sense that proper update with new knowledge is practically impossible: while updating the properties of one target species may improve things locally, the expense paid is rather steep, because such an update inherently increases the overall inconsistency of the tables. In addition, the sequential approach generally results in values (and their derived uncertainties) that reflect only local relationships inherent in the original measurable quantities. Lack of corrective feedback inherent in the sequential "adopt and

freeze” approach results in cumulative errors that keep worsening as more species are added to the table.

Our goal is to improve this situation by delivering an innovative algorithm, called as *Active Thermochemical Tables* (ATcT), that dynamically evolves through an advanced Grid-enabled collaborative environment, called *Open Collaborative Grid Services Architecture* (OCGSA). Novel modalities of deriving new scientific results can be stimulated by enabling a collaborative environment in which scientists can publish and share their results with others, perform sophisticated calculations that are individually not feasible, and integrate newly developed algorithms into a comprehensive framework. A detailed analysis of deploying ATcT on the Grid is provided in [18]; and the description of the OCGSA framework was presented in [1]. However, this paper concentrates on staging the ATcT application in a collaborative Grid environment, thereby merging the two symbiotic infrastructures.

The motivation for the work presented in this paper is two-fold. On one hand we enhance the functionality of ATcT by deploying it on a collaborative Grid infrastructure; and on the other hand we validate the collaborative Grid framework for the scientific problem solving process. The rest of this paper is organized as follows. An overview of the traditional mechanisms for deriving thermochemistry data is provided in Section 2. Section 3 describes the new network-oriented approach to optimize and improve on the existing methods. Section 4 justifies the need for using Grids in a distributed resource sharing environment. Section 5 discusses the role of computational Grids in collaborative problem solving process. Section 6 presents the extensions to the existing Grid architecture to enable ad-hoc collaborations. Finally, Section 7 summarizes the paper giving the scope of ongoing and future research.

## 2 Thermochemical Tables

Elementary to the discipline of thermochemistry is *enthalpy* ( $\Delta H_f^\circ$ ), which refers to the value of energy of a system when it is at constant pressure. The enthalpy relationships involved in

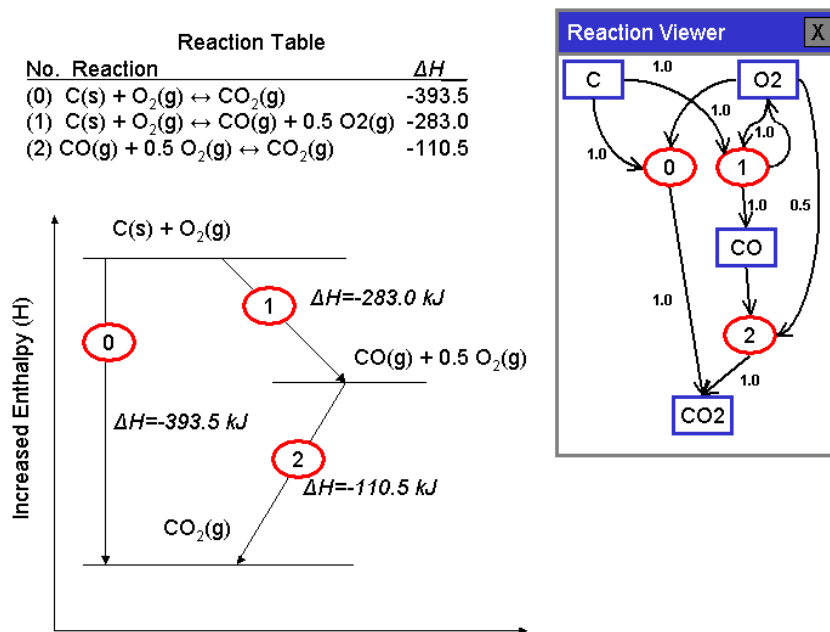


Figure 1: Enthalpy Diagrams and Thermochemical Reaction Tables

examining thermochemical equations are easily visualized by means of enthalpy diagrams, such as that shown on the lower-left hand corner in Figure 1.

In this diagram the equations are expressed as a graph with horizontal lines representing different values of the enthalpy. Typically the differences between these values are determined experimentally or can be derived by using thermodynamic laws from other enthalpy values. Values obtained from experiments, however, may contain errors (not shown in the diagram). Changes in the enthalpy are visualized by the distance between the lines. Based on the changes performed, different intermediate states (chemical species) may occur during the transition from one to the other final state. Thus, it is natural to visualize the transition with directed edges between the states. An alternative graph is displayed on the upper right-hand corner that emphasizes the possible states in which a chemical species (i.e., an ensemble of chemically identical molecular entities that can explore the same set of molecular energy levels on the time scale) can occur.

In traditional thermochemical tables the enthalpies of formation ( $\Delta H_f^\circ$ ) are developed with the help of an elaborate *sequential process*. In each step a new species is added while

all  $\Delta H_f^o$  values determined in previous steps are frozen. The enthalpy of the new species is determined at one temperature  $T$  from interconnecting measurements that are limited to those species already defined in previous steps. The temperature-dependent functions,  $C_p^o$ ,  $S^o$ ,  $(H_T^o - H_o^o)$ ,  $\Delta H_{fT}^o$ ,  $\Delta G_T^o$ , are developed by determining the partition function  $Q$ ,  $\ln Q$ ,  $T\partial(\ln Q)/\partial T$ , and  $T^2\partial^2(\ln Q)/\partial T^2$  from the species-specified quantities and the newly selected single temperature enthalpy [4].

The sequential process follows a *standard* order of chemical elements: O  $\rightarrow$  H  $\rightarrow$  halogens  $\rightarrow$  noble gases  $\rightarrow$  chalcogens  $\rightarrow$  pnictogens  $\rightarrow$  carbon period  $\rightarrow$  etc. For every chemical element introduced, the sequence starts at the standard state for that element, for which  $\Delta H_f^o = 0$  by convention. However, enthalpies of formation have complex hidden dependencies. These dependencies are backwards traceable, albeit with considerable manual effort, and in practice are not forwards traceable at all.

This sequential approach of introducing new enthalpies has several disadvantages. In particular, it results in  $\Delta H_f^o$  and error bars that do not properly reflect the global relationships implied by the species-interconnecting data used. The values and error bars reflect, at best, only local relationships to nearest neighbors. A cumulative error is introduced based on the frozen enthalpies in previous steps. Furthermore, the hidden relationships in conventional tables produce thermodynamic tables that are static in nature. Proper update with new knowledge is nearly impossible because forward relationships are nontransparent; hence, updating one species may improve the quality locally but increase the global inconsistencies. We present a new approach, based on active tables, that circumvents these disadvantages.

### 3 Active Thermochemical Tables

The active table approach treats the information in the species-interdependent data from the viewpoint of a thermochemical network [4, 13]. Every vertex (node) of this network represents the enthalpy of formation of one species. The species-interdependent data define

the topology of the network graph by providing the edges (links) in the network. Competing measurements provide multiple (parallel) links in the network. The relational information define the topology maps onto a set of equations, with the enthalpies of the involved species as unknowns, the stoichiometry of the reactions defining the coefficients, and the measurements and their error bars providing the free elements. Since the number of equations regularly exceeds the number of unknowns, the system is overdetermined. The best solution to the network is obtained in two steps.

The first step statistically analyzes the network with the goal of checking the error bars of individual data items for consistency. All inconsistencies are identified and proposals for their resolutions are generated. The resolutions are generally carried on through incrementing the error bars of the offending data items to their statistically indicated values, and the analysis is repeated. During subsequent steps the error bars of the offending data items usually oscillate up and down until self-consistency across the whole network is reached. The analysis is carried out in both unweighted and error-weighted space.

The second step occurs once the values and error bars of data items achieve consistency across the whole network. This step consists of finding simultaneously the optimal solution to all nodes by  $\chi^2$  minimization (in error-weighted space). The species-specific data is used to prepare the network for analysis by re-expressing all data items as reaction enthalpies at one common temperature. Once the optimal solution to the network has been found, the species-specific data is used to calculate the partition functions and hence develop the temperature dependence of the standard thermochemical functions. We term this approach *Active Thermochemical Tables* (ATcT).

It must be noted that, as opposed to the simultaneous solution described above, in the traditional sequential approach the nodes are solved one at a time in a prescribed sequence. Each of the steps corresponds to selecting a particular path in the network leading from solved nodes to the next node, and ignoring all other possible paths. Both the fact that an overdetermined system of coupled equations is solved for one unknown at the time and that

some of the equations are ignored contribute to the lack of global consistency in traditional tables.

Unlike the quantities found in a traditional table, the thermochemical quantities (and their error bars) obtained from the Active Table properly reflect the globalism of relationships implied by the underlying thermochemical network. All values and error bars are consistent in a global sense. An active table allows rapid update with new measurements (and possibly calculations) by globally propagating the new information through the table. Quality and integrity of the table are protected throughout the updates by error analysis, which runs in both directions: the error bar of the new experiment may shrink error bars in the table; however, the error bars of other experiments in the table might challenge, by means of the statistical analysis discussed above, the error bar assigned to the new experiment. In addition, an active table allows *what-if* scenarios and tests. Such tests provide a critical evaluation of the tested data and its impact on prior thermochemical knowledge or, if the tests correspond to a new experiment, provide feedback on sensitivity of the network to various measurements. The approach has also potential to become an interesting learning and education tool. An active table can provide a ranked list of links that are missing or are weak from a statistical viewpoint; in particular, it can provide pointers to the most useful new experiments or calculations to be conducted.

## 4 Grids for ATcT

Thermochemical data is available in a large number of databases that are generated electronically or published in academic papers. To increase the accuracy of the ATcT and its applicability in other scientific disciplines it is necessary to gather a database that is complete in the number of species and accurate based on state-of-the art experiments and deductions. Based on the number of publications and different experimental conditions it will be impossible by a single scientist to create such a database. Instead a new methodology is searched

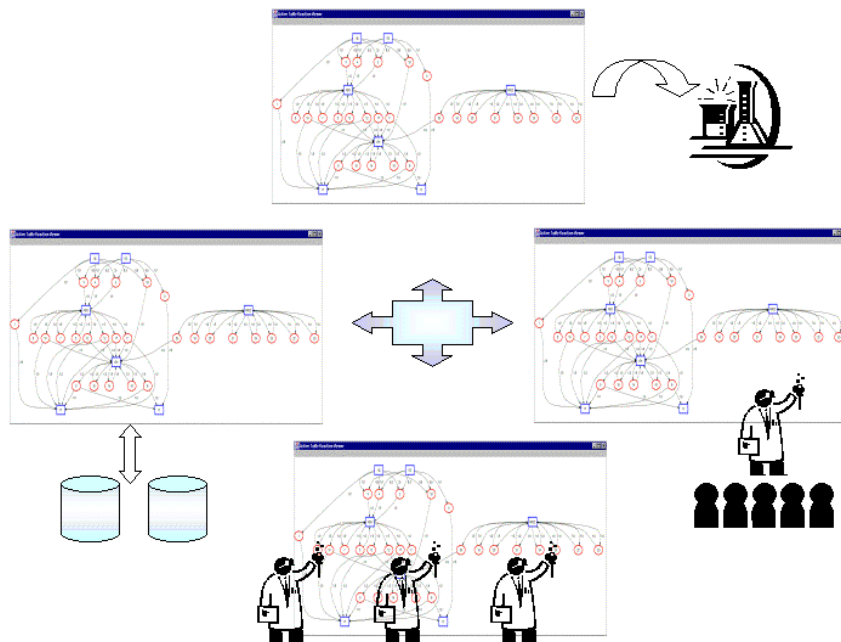


Figure 2: Collaboration through Active Murals

that enables contributions by a multitude of researchers. For example, consider a research collaboration for that has a suite of advanced scientific applications, including ATcT (See Figure 2). Researchers can invoke ATcT individually or collaborate in groups, exchanging their thermochemical data sets, interactively controlling their input to the active tables, and participating in collaborative discussions based on the results, thereby interactively validating the output. The new thermochemical data gathered by the community can be later used by individuals and groups. This enables an architecture that incrementally optimizes the thermochemical data based on community input.

To enable such a new modality of communication we need an infrastructure that is conducive to support collaborative interactions. This environment must fulfill the following important requirements:

**Collaboration:** the environment must support the interaction among scientists in geographically dispersed locations.

**Security:** the environment must protect from the loss of intellectual property and allows



restricted access to the data and compute resources.

**Standardization:** the environment must enable the scientist to use the standardized tools without the requirement to learn new technologies.

**Adaptivity:** the environment must be flexible enough to adapt to future changes in hardware and software.

**Extensibility:** the environment must allow the extension the current framework to include new capabilities and services based on application specific requirements.

These requirements are shared with many other scientific applications. Although several collaborative frameworks [6, 5] can be applied to the ATcT, this paper concentrates on the service-oriented Grid architecture to enable interactive problem solving.

## 5 Grid Computing

The term Grid computing [7] is commonly used to refer to a distributed infrastructure that promotes large-scale resource sharing in a dynamic multi-institutional virtual organization. Just as the electric power grid provides pervasive access to electric power, the computational Grid provides ubiquitous access to a large collection of compute-related resources and services. The Grid architecture identifies and solves a wide range of problems that arise as a result of large scale resource sharing among separately administered institutions. Some of these problems and solutions address important issues such as security, resource management, resource discovery, and data management. Recently, the Grid architecture has made a noticeable shift toward a services-oriented framework. A Web service is a platform-independent software component that is described by using a service description language and provides functionalities to publish its interfaces to some directory or registry of services. These published services can be discovered by querying the service directories or registries and can be invoked by using appropriate binding protocols. The services-oriented concept lends itself to an architecture that requires interoperability across heterogeneous platforms

and exposes all or part of its applications on different machines, platforms, and domains. Hence, the Open Grid Services Architecture (OGSA) [8, 11] was developed to provide an extensible framework of services that support Grid functionalities by combining the realms of Web services and Grid technologies.

Web services address the discovery and invocation of persistent services, whereas Grid technologies need to tackle issues related to creation, discovery, and destruction of transient services. Therefore, the OGSA introduced the notion of Grid service: a (potentially transient) Web service that provides a set of well-defined interfaces and that follows some standard conventions. OGSA provides a framework that supports

**instance creation** through Grid service factories that can be requested for new reliable transient service instances. OGSA provides a mechanism to individually identify these instances by associating state information with each service instance.

**lifetime management** to build transient service instances that can negotiate the duration and lifetime behavior with the service factory at the time of their creation. It may support soft-state lifetime management, thereby avoiding the need for explicit destruction of service instances.

**registration and discovery** to allow the registration and discovery of transient service instances.

**notification** to establish a publish/subscribe style of notification mechanism between service providers and service requesters, thereby supporting asynchronous peer-peer style of communication.

**security** to handle the required authentication and authorization policies defined by different application and domain administrators.

Sophisticated applications like the ATcT that can be controlled and steered by multiple users in an interactive fashion can benefit from the functionalities provided by OGSA. Although OGSA provides all the constructs (core services) to implement such platform-independent research collaboratories, considerable development effort is involved in each

case. Every application requires certain standard components at a minimum. Rather than developing these components in every project, it is worthwhile to invest some effort in identifying a common infrastructure that provides all the basic components required for collaborative applications and can be inherited by any stand-alone application with minimum effort to transform itself into a collaborative Grid service. We propose a framework that builds on top of the OGSA infrastructure, providing a set of services that can be used by any collaborative application including the ATcT. We call this new framework *Open Collaborative Grid Services Architecture* (OCGSA) [1, 3, 2].

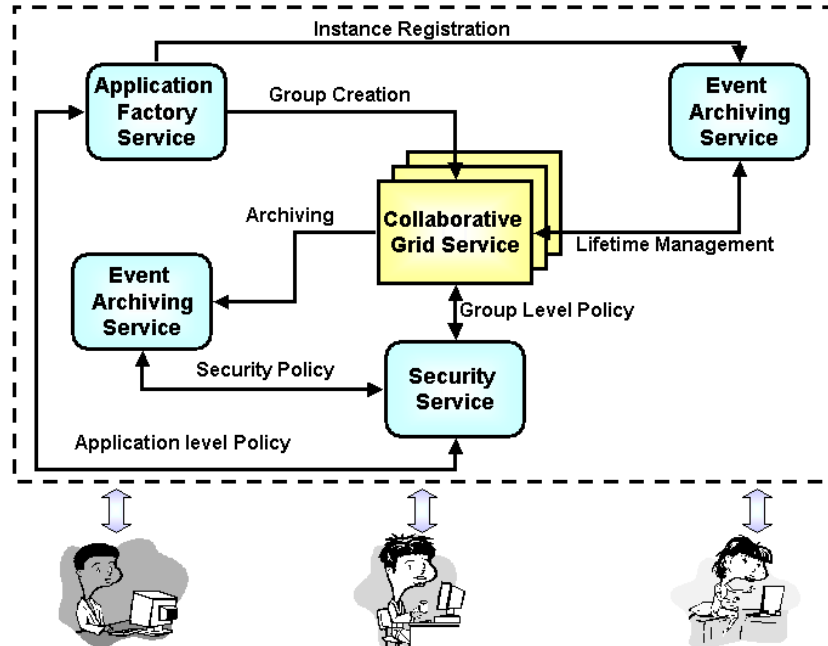


Figure 3: Open Collaborative Grid Services Architecture

## 6 OCGSA

The Open Collaborative Grid Services Architecture (OCGSA) composes a set of common components that can be easily customized for individual applications. The OCGSA framework enable users to form ad-hoc collaborative groups by interacting over a set of predefined notification topics. It provides appropriate lifetime management for individual groups, offers

an advanced discovery mechanism for service instances, establishes sophisticated security mechanisms at different levels of the application, and defines a pattern of communication between these components. Figure 3 shows the OCGSA framework depicting the common OGSA components for collaborative applications.

## 6.1 Collaborative Grid Service

The OGSA specification defines a Grid service as any Web service that conforms to a predefined interface. In other words, a Web service is a Grid service if it implements the GridService portType [8, 11]. The GridService portType provides a base interface for the service instance to act as a Grid service: providing maintenance of its metadata, lifetime management, and support for sophisticated security. OGSA allows Grid service developers to define customized metadata for individual applications in addition to the standard set of required service data elements. In a collaborative environment every Grid service instance warrants certain additional functionality and metadata definitions for the maintenance and discovery of individual service instances. In a collaborative context, this service instance is controlled and steered by a group of individuals, hence the service instance is referred to as a group instance (or simply group). Rather than redefining this metadata information for every collaborative Grid service, we introduce the notion of a collaborative Grid service: a Grid service with a set of predefined metadata elements and behaviors. In analogy with the GridService portType, we define a CollaborativeGridService portType that defines all the base interfaces that are required for a collaborative Grid service instance. Among other things, a CollaborativeGridService portType defines metadata regarding the creator of a group, name of the group, description of the group, current members of the group, and level of event archiving desired. It also provides the interface to handle security policies at different levels of the application (application level and group level), community interaction mechanisms, an adequate presence management for users, and capabilities for interactively controlling and steering the application.

A collaborative Grid service implements the GridService portType and the NotificationSource portType. This allows the collaborative service to generate publish/subscribe style notification messages to its clients by means of which they can interactively communicate. The NotificationSource portType is provided by OGSA; however, the OCGSA framework provides constructs in the collaborative Grid service descriptive language (cgddl) by which a developer is able to predefine a set of notification topics that will be supported by the collaborative Grid service. The collaborative Grid service description language is an extension to the Grid service description language (gsdl) defined by the OGSA and is used to describe details of a collaborative Grid service. The code generator will generate the necessary stubs for this definition, and the application-specific details for each notification topic can be implemented by the developer.

## 6.2 Security Service

Collaborative Grid services may be accessed by different groups of people from different institutions and domains. Trust levels within such groups are not as high as in a single organization. Most of the security constructs such as authentication, delegation of rights, confidentiality, and message integrity have been specified in the OGSA security draft. Nevertheless, the area where OCGSA security must build over the OGSA security is access control (authorization). The OCGSA framework specifies two levels of authorization. On the first level, called the application level, the application service provider dictates the policy about who can access the service based on the identity of the requester and the nature of the request. Subscription to notification topics and access to the event archiving services are based on this policy. A part of this policy decides who can create and register a new group. The second level is the group level access control. The group creator specifies who can join the group and the privileges of individual users within that group. For example, the group creator might authorize some users to only listen to the communication and not actively participate. The challenging aspect of OCGSA security is the ability to define and

enforce such policies dynamically at runtime without any prior configuration support from a specific system administrator.

The community authorization service (CAS) [12] of the Globus Toolkit [9] provides similar features to those desired by the OCGSA security service. Nevertheless, CAS requires that the security policies be defined statically and is maintained by a CAS administrator. The OCGSA security service provides an extension to the functionality offered by the CAS by specifying the policies dynamically. Hence, any user that desires to join a collaborative group communicates with the OCGSA security service by submitting his credentials in the form of X509 digital certificate. The security service then maps the user certificate to the application and group-level policies and accordingly rejects or grants the user a delegated proxy X509 certificate. The user uses this proxy certificate to communicate to the collaborative service instance where his proxy credentials are mapped against group access rights.

### **6.3 Registration and Discovery**

At the heart of any service-based architecture is the concept of service registration and discovery. The OCGSA framework provides constructs to formulate a registration service that establishes and maintains meta information about all the services registered with it. Using sophisticated querying mechanisms clients can discover the registered services from the information contained in the registry meta-data. The Registration service also provides a web service inspection (WS-Inspection) document that contains information about the various services registered with it. The OCGSA framework requires that every collaborative service be associated with a registration service. A single registration service can register instances of different collaborative services. The registration service used in OCGSA is a registration Grid service capable of supporting XPath/XQuery queries. Support for such a sophisticated querying mechanism is essential because OCGSA enables arbitrary users with appropriate credentials to query the registration service based on complex criteria and identify the service instance (group) that they would like to join from the set of services

returned by the query. Using a query language to locate services enhances the ability to discover Grid services in the OCGSA framework: services can be discovered not only by service names but by a variety of other meta information. Although a single registration service should suffice, the OCGSA framework is capable of supporting a set of registries, organized in a hierarchical or peer-to-peer fashion where a query submitted to any registry in the framework can be relayed to other registries if all the information is not present locally. The registration service must constantly communicate with the individual instances to maintain the lifetime management information up to date for the respective services and must automatically clean the entries for defunct groups.

## 6.4 Event Archiving Service

Another essential service for a collaborative framework is an event archiving service, which logs the messages or events communicated between online users of a group instance into a persistent database. Users that join a particular group at a later time can retrieve these events, thereby updating themselves about the group's prior activities. The event archiving mechanism also enables the group members to insert checkpoints into the event database and replay the entire or sections of the archived collaborative activity when the group is off-line. In the OCGSA framework, we assume an event to be a  $\langle \text{time, message} \rangle$  tuple. Message archiving can be implemented by using technologies such as Java message service (JMS) [10] or XML databases (XML:DB) [20]. However, OCGSA offers a sophisticated archiving service that functions more than a simple event database. The event archiving service in the OCGSA framework is a Grid service, efficiently handling the lifetime management policies of individual archives that can be set and manipulated by creators of corresponding group instances. Further, the event archiving service will implement the same security policies as implemented by the collaborative Grid services, hence forcing appropriate authorization policies at the application level and individual group level (See Section 6.2). This ensures that a malicious user lacking the proper credentials, who was denied access to an active group

earlier cannot have access to the communication archives when the group is inactive. Similar to discovering a collaborative Grid service, it is extremely important to efficiently and conveniently discover event archives. On fundamental level, the archiving system must support QueryByName style of information retrieval indicating if an archive for a particular group exists. Nevertheless, support must also be built to enable advanced queries of meta-data values of individual group archives similar to that of the OCGSA registration component. Hence, end users will be able to discover archives hosted by a set of archiving services based on complex queries regarding group names, lifetime validity, security authorization, and similar criteria.

In order to implement the functionality mentioned above, OCGSA supports an event archiving Grid service that provides an interface to XML:DB [20]. Hence, the Grid service can take care of the security and lifetime management issues, and the XML:DB (NXD, native XML database) provides the functionality for the data storage and retrieval using XPath queries. The XML:DB can be embedded within the Grid service, or can be located on a remote machine. The event archiving service can be considered as a specialization of a generic data storage Grid service that provides a Grid service interface to a native XML database (NXD) implementing all the generic operations required to manipulate the XML database that can store any XML document including event logs.

## 7 Ongoing Research

This paper discusses a novel approach in thermochemistry to optimize thermochemistry data utilizing the concepts of chemical networks, called as Active Thermochemical Tables (ATcT). It also enhances the applicability of ATcT by exploiting the collaborative features of the Grid. The developed framework allows for ad-hoc collaboration and interactive steering of the ATcT in a secure multi-institutional environment. Ongoing research on the proposed framework is concentrating on two complimentary areas: chemical sciences and computer



sciences.

From the chemical science perspective, efforts are being made to extend the concepts of ATcT into a more generic Active Tables. By conceptually expanding the Active Thermochemical Tables idea, one can envision adding similar "intelligence" to databases at other key levels of chemical scale. For example, Active Kinetic Rate Tables placed above the ATcT would make use of the relationships implied in a tabulation of elementary kinetic rate measurements and would be able to interact with ATcT, both as a user of information and as a contributor of new information. Above the Active Kinetic Rate Tables one may imagine an Active Mechanism Tables, followed by Active Reduced Mechanism Tables, etc., while down at the bottom there is room, at least in principle, for an Active Tables of Natural Constants, reproducing the periodic re-evaluations of fundamental physical constants. Such a series of stacked interacting Active Tables, would present a new paradigm to how science is conducted, because they would automatically exchange information up and down the chain, synergistically expanding the benefits through all involved scales. Such an arrangement would add a new dimension to the flow of information between scales and disciplines, allowing new knowledge to rapidly propagate through all involved chemical scales, and it would allow global consistency checks bridging several scales. In addition, such a development would provide seamless access to data. For example, a scientist would not have to worry about details of data at other scales, since the database at his scale of interest would automatically fetch the relevant data at other scales through interaction with other Active Tables. Finally, interacting Active Tables would allow global *what if* explorations and *weakest links* searches.

In order to support such a modality of communication, the middleware infrastructure for Grids is being extended to support collaborations for multi-disciplinary sciences. Promising technologies in Grid workflows is being investigated to develop an adaptable orchestration framework. Such a framework will enable support for complex inter-service communication patterns. The OCGSA framework will further be extended to encompass the following

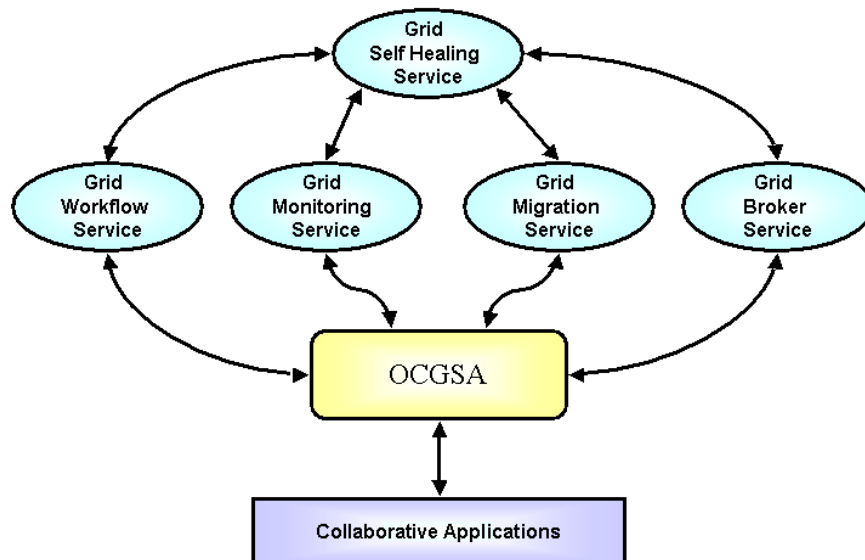


Figure 4: Advanced OCGSA Services

services (See Figure 4):

**Grid Broker Service** to deal with large numbers of calculations that are involved with future large-scale reactions and their real-time requirement for allowing interactive use [14].

**Grid Workflow Service** to enable the interplay of Grid services through workflow descriptions [15, 16].

**Grid Monitoring Service** to monitor the state of the hosting environment so that feedback to Grid services is provided, enabling the environment to react to state changes [17].

**Grid Migration Service** to be able to migrate services and jobs executed with a Grid Execution Factory Service to a location that is better suited, based on performance and quality of service descriptions and policies [19].

**Grid Self-Healing Service** to determine how and when it is necessary to change the dynamically instantiated Grid workflow applications (including preventive measurements

such as service replication, service migration, service checkpointing, and service monitoring).

## Acknowledgments

This work was supported by the Division of Chemical Sciences, Geosciences and Biosciences, Office of Basic Energy Sciences, and by the Mathematical, Information, and Computational Science Division subprogram of the Office of Advanced Scientific Computing Research, U.S. Department of Energy, under Contract W-31-109-Eng-38. Funding for the Globus Project research and development is provided by DARPA, DOE, and NSF. The work is performed with help of the SciDAC Collaboratory for Multi-scale Chemical Sciences (CMCS) and the SciDAC Java Commodity Grid project sponsored by the Department of Energy. We thank Jim Myers, Larry Rahn, David Leahy, David Montoya, Karen Schuchardt, Carmen Pancerella, Christine Yang, Ian Foster, the other members of the SciDAC CoG Kit project, the SciDAC CMCS project, and the Globus Project. The Globus Toolkit and Globus Project are trademarks held by the University of Chicago.

## References

- [1] Kaizar Amin, Sandeep Nijsure, and Gregor Laszewski. A Grid Services Framework for Collaborative Applications. In *SC'2002*, Baltimore, MD, 11-16 November 2002. (Poster).
- [2] Kaizar Amin, Sandeep Nijsure, and Gregor Laszewski. Collaborative Grid Services Framework. In *Globus World*, San Diego, CA, 17 January 2003. (Poster).
- [3] Kaizar Amin, Sandeep Nijsure, and Gregor von Laszewski. Open Collaborative Grid Services Architecture (OCGSA). In *Euroweb 2002 Conference, The Web and the GRID: from e-Science to e-Business*, pages 101–107, St Anne’s College Oxford, UK, 17-18 December 2002. The British Computer Society.
- [4] B.Ruscic, J. V. Michael, P. C. Redfern, L. A. Curtiss, and K. Raghavachari. Simultaneous adjustment of experimentally based enthalpies of formation of cf<sub>3</sub>x, x = nil, h, cl, br, i, cf<sub>3</sub>, cn, and a probe of g<sub>3</sub> theory. *J. Phys. Chem. A.*, 102:10889–10899, 1998.
- [5] CompreHensive collaborativE Framework (Chef). [www.chefproject.org/](http://www.chefproject.org/).

- [6] Collaboratory for Multi-scale Chemical Sciences (CMCS). <http://cmcs.ca.sandia.gov/>.
- [7] Ian Foster, Carl Kesselman, and Steve Tuecke. The anatomy of the grid: Enabling scalable virtual organizations. *International Journal of Supercomputer Applications*, 15(3), 2001.
- [8] Global Grid Forum (GGF). <http://www.gridforum.org/>.
- [9] The globus project. <http://www.globus.org>.
- [10] Java Message Service (JMS). <http://java.sun.com/products/jms/>.
- [11] Open Grid Services Architecture (OGSA). <http://www.globus.org/ogsa>.
- [12] Laura Pearlman, Von Welch, Ian Foster, Carl Kesselman, and Steve Tuecke. A community authorization service for group collaboration. In *IEEE Workshop on Policies for Distributed Systems and Networks*, 2002.
- [13] B. Ruscic, M. Litorja, and R. L. Asher. Ionization energy of methylene revisited: Improved values for the enthalpy of formation of  $\text{CH}_2$  and the bond energy of  $\text{CH}_3$  via simultaneous solution of the local thermochemical network. *J. Phys. Chem. A.*, 103:8625–8633, 1999.
- [14] Gregor von Laszewski and Ian Foster. Grid Infrastructure to Support Science Portals for Large Scale Instruments. In *Proceedings of the Workshop Distributed Computing on the Web (DCW)*, pages 1–16. University of Rostock, Germany, 21–23 June 1999. (*Invited Talk*).
- [15] Gregor von Laszewski, Ian Foster, Jarek Gawor, and Peter Lane. A Java Commodity Grid Kit. *Concurrency and Computation: Practice and Experience*, 13(8-9):643–662, 2001.
- [16] Gregor von Laszewski, Ian Foster, Jarek Gawor, Peter Lane, Nell Rehn, and Mike Russell. Designing Grid-based Problem Solving Environments and Portals. In *34th Hawaiian International Conference on System Science*, Maui, Hawaii, 3–6 January 2001. <http://www.mcs.anl.gov/~laszewsk/papers/cog-pse-final.pdf>, <http://computer.org/Proceedings/hicss/0981/volume%209/0981toc.htm>.
- [17] Gregor von Laszewski, Jarek Gawor, Carlos J. Peña, and Ian Foster. InfoGram: A Peer-to-Peer Information and Job Submission Service. In *Proceedings of the 11th Symposium on High Performance Distributed Computing*, pages 333–342, Edinbrough, U.K., 24–26 July 2002.
- [18] Gregor von Laszewski, Branko Ruscic, Patrick Wagstrom, Sriram Krishnan, Kaizar Amin, Sandeep Nijssure, Reinhardt Pinzon, Melita L. Morton, Sandra Bittner, Mike Minkoff, Al Wagner, and John C. Hewson. A Grid Service Based Active Thermochemical Table Framework. In *Third International Workshop on Grid Computing*, Lecture Notes in Computer Science, pages 25–38, Baltimore, MD, 18 November 2002. Springer.
- [19] Gregor von Laszewski, Kazuyuki Shudo, and Yoichi Muraoka. Grid-based Asynchronous Migration of Execution Context in Java Virtual Machines. In Arndt Bode, Thomas Ludwig, Wolfgang Karl, and Roland Wismüller, editors, *Proceedings of EuroPar 2000*, volume 1900 of *Lecture Notes in Computer Science*, pages 22–34, Munich, Germany, 29 August - 1 September 2000. Springer. (*Invited Talk*).
- [20] The XML:DB Initiative for XML Databases. [www.xmldb.org](http://www.xmldb.org).